Polar Meteorol. Glaciol., **12**, 97–103, 1998

CLOUD EXTRACTION FROM POLAR SATELLITE DATA USING MODIFIED MAHALANOBIS CLASSIFIER

Mamoru KUBO[1], Hideo SAITO[1], Ken'ichiro MURAMOTO[1] and Takashi YAMANOUCHI[2]

[1]*Faculty of Engineering, Kanazawa University, Kodatsuno 2-chome, Kanazawa 920-8667*
[2]*National Institute of Polar Research, Kaga 1-chome, Itabashi-ku, Tokyo 173-8515*

Abstract: In the polar region, it is difficult to discriminate between clouds and ground surface from satellite visible or infrared data, because of the high albedo and low surface temperature of snow and ice cover. In addition, since the latitude is high, visible channels cannot be used during winter. In this paper, a method to extract clouds using only NOAA/AVHRR channel 4 is proposed. Using geographical information, the AVHRR image was first segmented into two regions: sea and land. Afterward, cloud extraction was performed for each region separately by minimum distance classifier using image features. To improve the error rate of the classification, we apply thresholds to the discriminant function used by the Mahalanobis classifier.

1. Introduction

Clouds have a major role in the radiative processes in planetary atmospheres both in the absorption and reflection of solar radiation and in the emission of thermal energy (CHAHINE, 1982; HARTMANN and SHORT, 1980; KEY and BARRY, 1989; SALTZMAN and MORITZ, 1980). A meteorological satellite provides a great deal of information about the surface of the earth. Several studies have attempted detecting cloud cover from visible and infrared satellite-measured radiance. However, in the polar region, cloud, snow and ice have almost the same albedo in the visible channel and the same brightness temperature in the infrared channel; therefore it is difficult to distinguish among these regions using only the threshold of gray level of a satellite image (COAKLEY and BRETHERTON, 1982; DESBOIS *et al.*, 1982). In addition, because of high latitude, visible channels cannot be used during winter. To classify the area from satellite images in all seasons, we have to use only an infrared channel.

YAMANOUCHI *et al.* (1987) reported a method to detect cloud using only the infrared channels of NOAA/AVHRR. This method used the difference between channels 3 and 4. However, since channel 3 contains both reflected solar radiation and terrestrial radiation, the characteristics of this channel differ in daytime and nighttime, and in summer and winter.

We have reported techniques for classifying Antarctic satellite images into cloud, sea ice and ground using only channel 4 data (MURAMOTO and YAMANOUCHI, 1996). That algorithm consisted of two major approaches: extraction of image features and minimum distance classification. However, misclassification was frequent. In this paper, an improved classification method which reduces the error rate is proposed.

2. Satellite Data Processing

Figure 1 shows the channel 4 image of NOAA/AVHRR on October 22, 1988, near Syowa Station. The spatial resolution is 2.2 km. The area is composed of 512 by 512 pixels covering a land area of 1100 by 1100 km. At each pixel location, the image brightness was quantified into 256 gray levels for computer graphics display. The lower part of the scene covers the continental snowfield and the upper part covers the sea area, which sometimes contains sea ice. To detect clouds in these images, we have proposed a method of classification using image features.

All the features presented in this work were obtained using template techniques. A subregion was defined as a 32 by 32 pixels block. The input pixels within the subregion were used to calculate an output pixel value. The subregion was then shifted one pixel to the right on the same line and the process was repeated.

3. Segmentation into Sea and Land Areas

The image features used in this work are:

(1) average of brightness temperature,
(2) standard deviation of brightness temperature,
(3) local fractal dimension,
(4) uniformity of texture,
(5) correlation of texture.

The fractal dimension is a quantitative measure of structural similarity (PENTLAND, 1984). The method relies on the assumption that regions of an image having a particular structure will usually produce a fractal gray level surface, with a particular value of the fractal dimension.

When all pixels were classified using these image features, the error rate was 36.7%, because of feature similarity among cloud, sea and ground (MURAMOTO *et al.*, 1998). Using geographical information, the AVHRR image was segmented into two regions: sea and land. The image features used for cloud extraction in each region are summarized in Table 1. Three features were computed for cloud extraction in the sea region and four features for the land region.

Table 1. Image features.

Land	Average of brightness temperature Local fractal dimension Correlation
Sea	Average of brightness temperature Standard deviation of brightness temperature Uniformity Correlation

4. Mahalanobis Classification

In order to classify clouds, sea ice and ground, a representative area for each desired class was selected subjectively using infrared imagery (channel 4). Figure 2 shows the estimated image. The whole area was classified manually using channel 1 and the difference between channels 3 and 4 to obtain supervised data for estimating the classification results. Each feature of image data in the selected areas was computed for training samples.

Supposing that x is the position of the pixel to be classified, the Mahalanobis distance is defined as

$$d_i=(x-m_i)\ \Sigma_i\ (x-m_i), \quad (1)$$

where m_i and Σ_i are the mean vector and covariance matrix of the data in class w_i, $i=1,\ldots, M$, determined from training data. Thus the mahalanobis distance retains a degree of direction sensitivity *via* the covariance matrix.

Mahalanobis classification is performed on the basis of

$$x \in w_i \text{ if } d_i < d_j \text{ for all } j \neq i. \quad (2)$$

Every pixel in the image was classified into one of the classes. Poorly classified pixels lie near the decision surface. To remove these pixels, thresholds are applied to the discriminant functions. The decision rule of eq. (2) was added:

$$|d_i-d_j| > d_{th}. \quad (3)$$

Using thresholds, a pixel can be left as unclassified.

5. Results and Discussion

Figure 3 shows the classification of the land region in Fig. 1. In Fig. 3a and 3b, pixels that are correctly classified as clouds and ground are colored in shades of white and dark gray. Pixels that are colored black are misclassified to clouds or ground.

Table 2. Classified pixels of land region (%).

Threshold (d_{th})	Unclassified	Cloud (error)	Ground (error)	Total error
—	—	16.8 (4.3)	83.1 (1.5)	5.9
10	8.5	12.7 (1.9)	78.9 (0.4)	2.2

Table 3. Classified pixels of sea region (%).

Threshold (d_{th})	Unclassified	Cloud (error)	Sea (error)	Total error
—	—	53.1 (5.2)	46.9 (9.8)	15.0
10	11.3	46.3 (4.1)	42.4 (6.8)	10.9

When all pixels in the land region were classified, the error rate was 5.9%, the correct reference being Fig. 2. The error rate was decreased to 2.2% by the application of thresholds, though some areas were left unclassified as shown in Table 2. Figure 3c shows the result of the application of thresholds in the land region. Pixels that are colored black are unclassified pixels.

Figure 4 shows a result of classification of the sea region. When all pixels in the sea region were classified, the error rate was 15.0% as shown in Table 3. The error rate was decreased from 15.0% to 10.9% by the application of thresholds, however

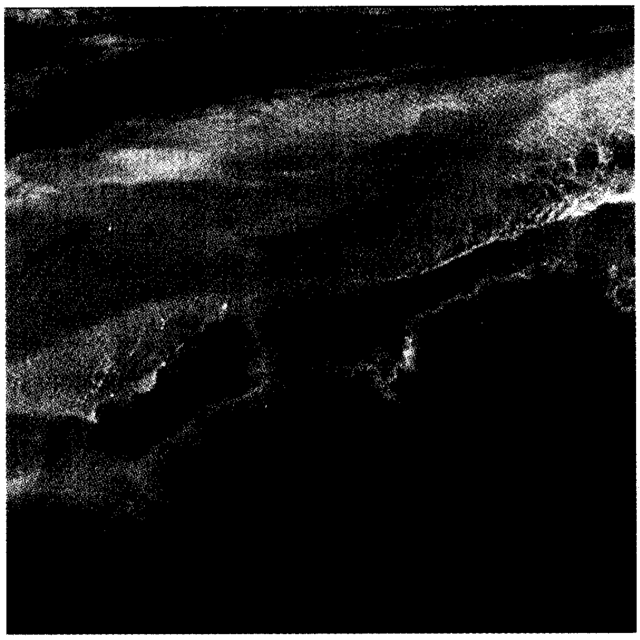

Fig. 1. NOAA/AVHRR channel 4 image.

Fig. 2. Estimation image.

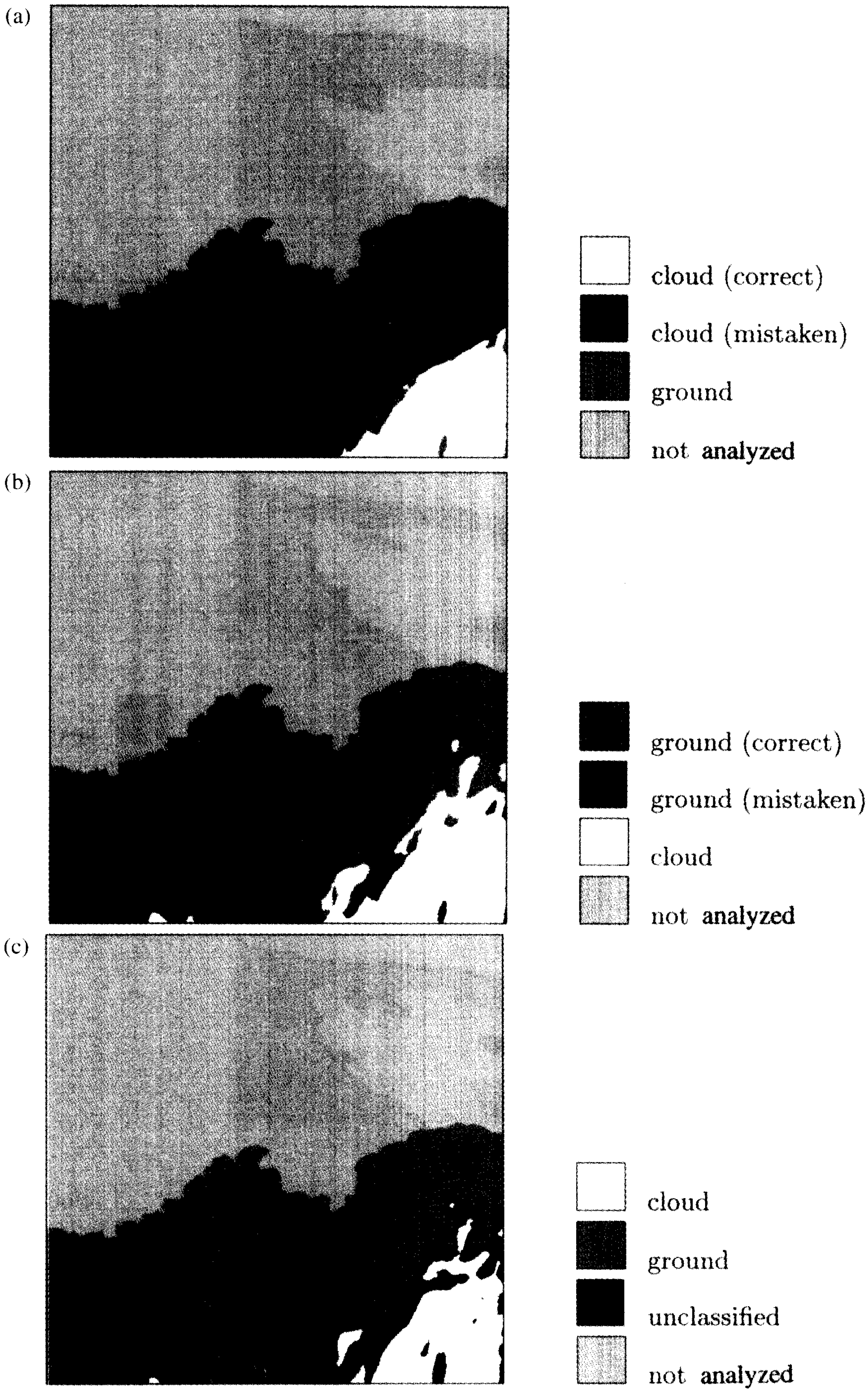

Fig. 3. *Classification of land region (lower part). (a) White is correctly classified as cloud. Black is misclassified as cloud. (b) Dark gray is correctly classified as ground. Black is misclassified as ground. (c) Applying a threshold to remove misclassification. Unclassified pixels are black. White pixels are classified as cloud. Dark gray is ground.*

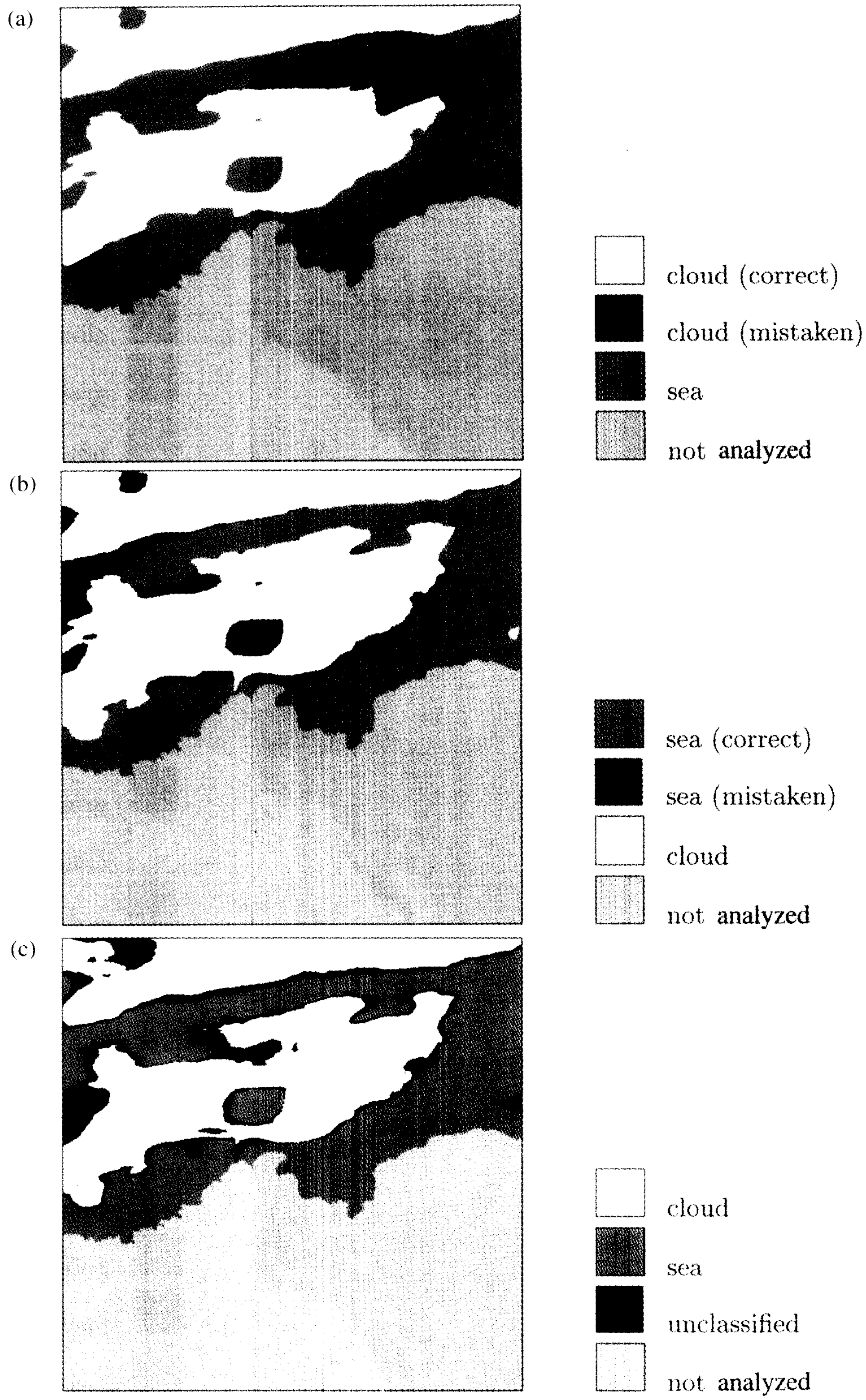

Fig. 4. Classification of sea region (upper part). (a) White is correctly classified as cloud. Black is misclassified as cloud. (b) Dark gray is correctly classified as sea. Black is misclassified as sea. (c) Applying a threshold to remove misclassification. Unclassified pixels are black. White pixels are classified as cloud. Dark gray is sea.

unclassified area increased up to 11.3%.

It is important that both accuracy and unclassified area are not optimized simultaneously; instead, a compromising threshold value has to be found. In this case, a value of 10 was used.

It is still unclear why the classifier mistakenly labeled some areas as other than cloud, and sometimes failed to distinguish clouds. A possible explanation is that since training data were taken from very distinct areas, and the mislabeled pixels resided in boundary regions, the classifier failed when cloud cover was thin or different.

6. Conclusions

Detection of Antarctic clouds is important because of their strong radiation influence on the energy balance of snow and ice surfaces. A method to extract cloud from an Antarctic satellite image using a single infrared channel is proposed. The average and standard deviation of brightness temperature, fractal dimension and textural features of the image data were used to classify cloud, sea and ground. When all pixels were classified, the error rate was large. To decrease the error rate of the classification, the image was segmented into two regions (sea and land) using geographical information. To remove misclassification, a threshold was applied to the discriminant functions. The error rate was significantly reduced using this technique.

References

CHAHINE, M. T. (1982): Remote sensing of cloud parameters. J. Atmos. Sci., **39**, 159–170.

COAKLEY, J. A. and BRETHERTON, F. P. (1982): Cloud cover from high-resolution scanner data: Detecting and allowing for partially filled fields of view. J. Geophys. Res., **87**, 4917–4932.

DESBOIS, M., SEZE, G. and SZEJWACH, G. (1982): Automatic classification of clouds on METEOSAT imagery: Application to high-level clouds. J. Appl. Meteorol., **21**, 401–412.

HARTMANN, D. L. and SHORT, D. A. (1980): On the use of Earth radiation budget statistics for studies of clouds and climate. J. Atmos. Sci., **37**, 1233–1250.

KEY, J. and BARRY, R.G. (1989): Cloud cover analysis with Arctic AVHRR data. J. Geophys. Res., **94**, 18521–18535.

MURAMOTO, K. and YAMANOUCHI, T. (1996): Classification of polar satellite data using image features and decision tree classifier. Proc. NIPR Symp. Polar Meteorol. Glaciol., **10**, 127–137.

MURAMOTO, K., KUBO, M., SAITO, H. and YAMANOUCHI, T. (1998): Classification of polar satellite data using minimum distance method. Mem. Natl Inst. Polar Res., Spec. Issue, **52**, 149–157.

PENTLAND, A. (1984): Fractal-based description of natural scenes. IEEE Trans. Pattern Anal. Machine Intell., **6**, 666–674.

SALTZMAN, B. and MORITZ, R. E. (1980): A time-dependent climatic feedback system involving sea-ice extent, ocean temperature, and CO_2. Tellus, **32**, 93–118.

YAMANOUCHI, T., SUZUKI, K. and KAWAGUCHI, S. (1987): Detection of clouds in Antarctica from infrared multispectral data of AVHRR. J. Meteorol. Soc. Jpn., **65**, 949–961.

(Received January 26, 1998; Revised manuscript accepted May 6, 1998)